

# Contrast in Multipath Interference and Quantum Coherence

Kai von Prillwitz,<sup>1</sup> Łukasz Rudnicki,<sup>1</sup> and Florian Mintert<sup>1,2</sup>

<sup>1</sup>Freiburg Institute for Advanced Studies, Albert-Ludwigs University of Freiburg, Albertstraße 19, 79104 Freiburg, Germany

<sup>2</sup>Department of Physics, Imperial College London, London SW7 2AZ, United Kingdom

We develop a rigorous connection between statistical properties of an interference pattern and the coherence properties of the underlying quantum state. With explicit examples, we demonstrate that even for inaccurate reconstructions of interference patterns properly defined statistical moments permit a reliable characterization of quantum coherence.

## I. INTRODUCTION

Interference resulting from quantum coherence causes an abundance of effects that contradict our classical intuition. Most people would probably be inclined to negate both the interference of independent photons [1] or of mesoscopic molecules [2, 3] if there was no clear experimental evidence for the existence of both effects; and despite the fact that interference phenomena in quantum mechanical systems have been observed for more than a century, we can still find large missing pieces in our understanding of quantum coherence.

The fact that a coherent superposition of at least two path alternatives (two state-vectors in a more general, abstract description) is necessary for an *interference pattern* to emerge, and that the achievable contrast increases with the number of states that are coherently over-imposed is one of the best established notions of elementary physics. However, going beyond this qualitative observation, our intuition is typically not able to answer the question of how many path-alternatives are needed to generate a particular interference pattern with reduced contrast. The overall aim of this paper is to explore the information content stored in the interference pattern, and to develop a framework which addresses the above question.

The formal definition of quantum coherence requires a set of mutually orthogonal states  $|j\rangle$ ,  $j = 1, \dots, d$  with respect to which coherence is defined. In an interferometric situation these states would correspond to different path alternatives and the number of paths that are being taken coherently is often referred to in terms of the lateral coherence length. In the case of molecular networks one is typically interested in the number of chromophores over which an excitation is coherently distributed, so that coherent delocalization is defined in terms of the excited states of the individual chromophores [5, 6]; and in transport theory or quantum thermodynamics this reference basis is given by energy eigenstates [7, 8].

In general, a pure state  $|\Psi\rangle$  is considered  $k$ -coherent in terms of a given set of basis states  $|j\rangle$ , if at least  $k$  of the amplitudes  $\langle\Psi|j\rangle$  are non-vanishing. Since decoherence processes which are unavoidably present in actual physical situations result in the deterioration of quantum coherence, the description in terms of mixed states or density matrices becomes necessary. As every mixed state  $\rho$  can intuitively be understood as an average over

pure states, averaging over *incoherent* states  $|\Upsilon_i\rangle$  will not result in any interference phenomena. Consequently, any mixed state  $\rho = \sum_i p_i |\Upsilon_i\rangle \langle\Upsilon_i|$  that can be decomposed into a mixture of incoherent states with  $p_i \geq 0$  is considered incoherent. Analogously, every mixed state that can be expressed in terms of an average over pure states with no more than  $k$ -coherence is *not*  $k+1$ -coherent. This motivates the commonly employed definition [9–11] that a mixed state  $\rho^{(k)}$  is  $k$ -coherent, if any ensemble  $\{|\Psi_i\rangle\}$  that satisfies  $\rho^{(k)} = \sum_i p_i |\Psi_i\rangle \langle\Psi_i|$  for some set of probabilities  $p_i$  contains at least one  $k$ -coherent state vector. The notion of  $k$ -coherence is thus similar to the concept of multipartite entanglement, since a mixed state is called  $k$ -partite entangled if any of its ensemble decompositions involves at least one  $k$ -partite entangled pure state [12].

Quantum coherence has recently been recognized as a resource [8, 9, 13–16] in the sense that there are processes whose realization is facilitated by the consumption of coherence. Various tools known from the entanglement theory have thus been adapted for the classification and quantification of quantum coherence [10, 13, 17, 18]. Reconstruction of the complete density matrix is required to assess most of these tools, and only few schemes work with fewer observables to be measured [9, 19]. On the one hand, this fact poses a rather high threshold for the analysis of coherence in laboratory experiments, and, on the other hand, the abstract nature of the aforementioned tools limits the intuition that might be gained from their use.

We will strive for the identification of quantum coherence based on the interference pattern only. If an interference pattern can be decomposed into a sum of patterns resulting from  $k$ -path interference, then, this pattern does not permit to conclude on  $k+1$ -coherence (as exemplarily depicted in Fig. 1 for  $d = 3$  and  $k = 2$ ). We will therefore identify (in Section II) properties of interference pattern asserting that such a decomposition is not possible. In Section III we provide numerical evidence underlining the performance of the tools developed in Section II. The numerical studies have been designed to capture major practical issues, such as difficulties with the proper identification of the maximum of a complicated interference pattern, or its coarse graining.

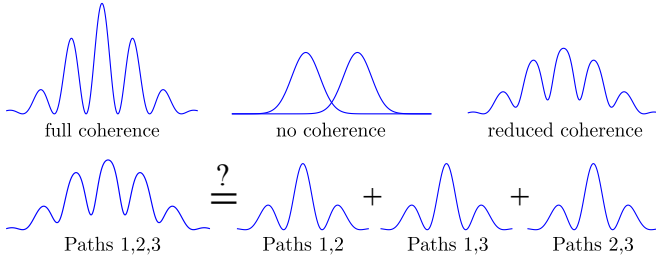


Figure 1. Upper part: interference patterns corresponding to different types of coherence in mixed-states. Lower part: an interference pattern that results from a state distributed over three path alternatives. Reduced coherence between the three paths can sometimes be expressed as the sum of two-path interference patterns. In that case the pattern can result from a two-coherent state.

## II. THE INTERFERENCE PATTERN AS A COHERENCE CLASSIFIER

We consider a rather general physical situation in which a superposition of different states is being established, and a certain level of decoherence results in the fact that this superposition is not perfectly coherent. A specific realization of such a situation would be a Mach-Zehnder type of interferometer, as schematically depicted in Fig. 2 for  $d = 3$ , where the different path-alternatives define the basis states  $|j\rangle$ . An incoming object impinges

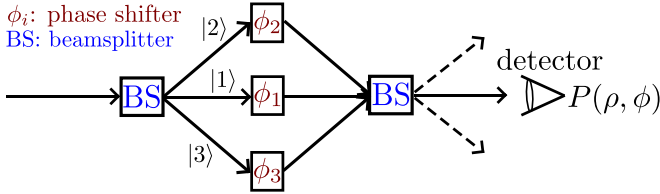


Figure 2. Three-path extension of the Mach-Zehnder interferometer setup. An incoming particle is splitted into three contributions. After passing the individual phase shifters the three beams are recombined at the second beam splitter where they interfere.

on a beam-splitter (BS), that creates a coherent superposition of the basis states. The phase shifters  $\phi_i$  permit to generate an interference pattern that can be read off, once the object has crossed the second beam splitter.

The interference pattern

$$P(\rho, \phi) = \langle \Phi | \rho | \Phi \rangle \equiv 1 + \sum_{j \neq m} \rho_{jm} e^{i(\phi_j - \phi_m)}, \quad (1)$$

with  $|\Phi\rangle = \sum_{j=1}^d e^{-i\phi_j} |j\rangle$  is defined as the normalized probability distribution to observe an object in the output mode, where  $\rho$  is the state before crossing the second beam splitter. Since on average only one out of  $d$  objects exits through the output mode, the interference pattern is given in terms of an over-normalized state with  $\langle \Phi | \Phi \rangle = d$ .

The simplest case of  $d = 2$  path alternatives corresponds to the original Mach-Zehnder-Interferometer, in which one can record the interference pattern by tuning a single phase shifter. In general, the interference pattern is obtained by tuning  $d - 1$  phase shifters. Beyond the obvious increase of dimensionality, also the structure of the pattern typically gets more complicated with growing  $d$  since the dependence of the detection probabilities on the phases  $\phi_i$  gets more sensitive. Our aim shall thus be to capture a more global part of the desired information, which is robust against small deviations of the tuned parameters. To this end we examine various *moments* of the interference pattern in question.

### A. Moments of the interference pattern

One can certainly obtain some information on coherence from the maximum of the interference pattern

$$\max_{\phi} P(\rho, \phi), \quad (2)$$

as its value, when larger than  $k - 1$ , unambiguously identifies  $\rho$  to be  $k$ -coherent. In practice, however, this is not necessarily the best choice. In particular, for highly coherent states, the interference pattern is a rapidly oscillating function so that optimizations will often identify only local maxima with a resulting under-estimation of coherence properties. Since, again, for highly coherent states, the optimum is given by a very narrow peak, an extremely accurate reconstruction of the interference pattern becomes necessary.

A much more practical alternative would be to employ the uniform statistical moments

$$m_n = \int_0^{2\pi} \frac{d^d \phi}{(2\pi)^d} P^n(\rho, \phi) \equiv \int_0^{2\pi} \frac{d^d \phi}{(2\pi)^d} \langle \Phi | \rho | \Phi \rangle^n. \quad (3)$$

The first moment  $m_1 = 1$  is just the norm of the interference pattern, but the higher moments carry non-trivial information. One would expect that increasing the order of moments improves the identification of  $k$ -coherence, because taking the limit  $\lim_{n \rightarrow \infty} (m_n)^{1/n}$  is equivalent to finding the maximum (2). On the other hand the required accuracy of the reconstructed pattern (*e.g.* from experimental data) necessary to assess a moment grows with  $n$  [20].

We thus strive for an approach that is based on moments of reasonably low order, which are more robust against small deviations of the interference pattern. In order to find a good compromise between a sensitive identification and robustness with respect to imperfections, we utilize the *generalized moments*

$$Q_n = \int d^d \phi F(\phi) P^n(\rho, \phi), \quad (4)$$

defined in terms of a suitably chosen  $d$ -dimensional probability distribution  $F(\phi)$ . The simplest case  $F(\phi) =$

$(2\pi)^{-d}$  reproduces the uniform moments Eq. (3). In the opposite case, when  $F(\phi)$  is strongly localized around the maximum of the interference pattern, the value of the generalized moment  $Q_n$  approximates the maximum  $\max_{\phi} P^n$ . This specific choice calls for the search of an optimal probability distribution which might be flawed by the same issues as encountered for Eq. (2). With a sufficiently wide distribution  $F$ , on the other hand, the optimization landscape is substantially flatter than in Eq. (2) what eases the optimization a lot. Using the generalized moments of low order together with the distribution  $F$  encoding additional information, namely the expected position of the maximum of the interference pattern, one can reasonably merge advantages of both interference peaks (Eq. (2)) and regular statistical moments (Eq. (3)), avoiding complications brought by each of two approaches alone.

### B. The wrapped normal distribution

Since the interference pattern defined in Eq. (1) is given in terms of all the phases  $\phi_j$  as independent variables, it is reasonable to define  $F(\phi)$  in terms of independent distributions for each phase, *i.e.*  $F(\phi) = \prod_{j=1}^d f(\phi_j; \mu_j)$ . The variable  $\phi_j$  is thus distributed according to  $f(\phi_j; \mu_j)$ , with  $f(\cdot)$  having the same functional form for all  $\phi_j$ . By  $\mu_j$  we denote the expectation value of each of these distributions. Although strictly speaking not necessary, we will assume equal width of all distributions and denote their standard deviation by  $\sigma$ .

An evaluation of Eq. (4) requires the construction of so-called *trigonometric moments*  $\Theta_n(\mu) \equiv \int_0^{2\pi} d\phi f(\phi; \mu) e^{in\phi}$  [21] defined for any integer  $n$ . Due to the fact, that the phases  $\phi_j$  are defined only in an interval of width of  $2\pi$ , this step can be done explicitly for most typically employed distributions like the Lorentz or Gauss distributions. To this end, it is helpful to take advantage of a *wrapped* version of a distribution [21]. In the case of the wrapped normal distribution, the trigonometric moments are equal to the characteristic function of the normal (unwrapped) distribution evaluated at integer arguments,  $\Theta_n(\mu) = e^{in\mu} R_n$ , where  $R_n = e^{-n^2\sigma^2/2}$ . With the help of the function  $\Theta_n(\mu)$  one can perform the integration in Eq. (4) and express the generalized moments as

$$Q_n = \sum_{i_1, i_2, \dots, i_{2n}=1}^d \rho_{i_1 i_{n+1}} \rho_{i_2 i_{n+2}} \dots \rho_{i_n i_{2n}} \prod_{j=1}^d \Theta_{n_j}(\mu_j), \quad (5)$$

$$\text{with } n_j = \sum_{l=1}^n \delta_{j, i_l} - \sum_{l=n+1}^{2n} \delta_{j, i_l}. \quad (6)$$

As argued above, the use of low order moments is desirable. We will therefore focus in the following on  $n = 1$ ,  $n = 2$  and  $n = 3$ . There is however no fundamental obstacle for generalizations to higher values of  $n$ .

### C. Threshold values

Before one can use the moments defined in Eqs. (4) or (5) to rigorously identify coherence properties, one needs to find the maximum that  $Q_n$  can adopt for  $k$ -coherent states. In the present case such an optimization can be done explicitly, confirming that the maximum among all  $k$ -coherent states is provided by

$$|W_k\rangle = \frac{1}{\sqrt{k}} \sum_{j=1}^k e^{-i\varphi_j} |j\rangle, \quad (7)$$

*i.e.* a perfectly balanced coherent superposition of  $k$  basis states.

To arrive at this conclusion, one may first realize that the generalized moments  $Q_n$  are convex, *i.e.*

$$Q_n(\eta\rho_1 + (1-\eta)\rho_2) \leq \eta Q_n(\rho_1) + (1-\eta)Q_n(\rho_2), \quad (8)$$

for  $0 \leq \eta \leq 1$  and any pair of density matrices  $\rho_1$  and  $\rho_2$ . This is a direct consequence of the two facts that the  $n$ th power of a linear functional like  $\langle \Phi | \rho | \Phi \rangle^n$  is convex for  $n \geq 1$ , and that the integral  $\int d^d\phi F(\phi)$  preserves convexity as  $F$  is non-negative. Since states  $\rho^{(k)}$  that are at most  $k$ -coherent (for any value of  $k$ ) define a convex set (*i.e.*  $\eta\rho_1^{(k)} + (1-\eta)\rho_2^{(k)}$  is no more than  $k$ -coherent) the maximum of  $Q_n$  over  $k$ -coherent density matrices is always reached for a pure state.

The most general  $k$ -coherent pure state reads

$$|\Psi^{(k)}\rangle = \sum_{j=1}^k \sqrt{\lambda_j} e^{-i\varphi_j} |j\rangle, \text{ with } \lambda_j \geq 0, \quad (9)$$

assuming (without loss of generality) that exactly the first  $k$  basis states are comprised with non-vanishing weights  $\lambda_j$  in the coherent superposition. In Appendix A it is shown that the optimization of the phase factors  $e^{-i\varphi_j}$  can be performed independently of the optimization over the real amplitudes  $\lambda_j$ , and that the maximum is obtained if the  $\varphi_j$  coincide with the expectation values  $\mu_j$  of  $f(\phi_j)$ . Also the remaining optimization over the  $\lambda_j$  can be performed very generally. As further shown in Appendix A, the quantity to be optimized is a Schur-concave function which is maximized for  $\lambda_j = 1/k$  for  $j = 1, \dots, k$ .

Summarizing the above considerations, the maximum of  $Q_n$  that can be adopted for  $k$ -coherent states with a given distribution  $F$  reads

$$Q_n^{(k)} = \max_{\rho^{(k)}} Q_n(\rho^{(k)}) = k^{1-n} \sum_{l=0}^{n^2} v_l^{(n,k)} e^{-l\sigma^2}, \quad (10)$$

with the coefficients  $v_l^{(n,k)}$  given in Table I. Any excess of this threshold value is an unambiguous identification of coherence properties beyond  $k$ -coherence.

$l$	$n = 1$	$n = 2$	$n = 3$
0	1	$2k - 1$	$4 - 9k + 6k^2$
1	$K_1$	4	$3K_1[11 + 3k(2k - 5)]$
2	-	$K_1K_2K_3$	$9K_1K_2^2K_3$
3	-	$2K_1K_2$	$K_1K_2^2(45 + kK_{10})$
4	-	$K_1$	$3K_1[k(55 + 2kK_9) - 52]$
5	-	-	$9K_1K_2K_3$
6	-	-	$2K_1K_2K_3$
7	-	-	$6K_1K_2$
8	-	-	0
9	-	-	$K_1$

Table I. The coefficients  $v_l^{(n,k)}$  that characterize the threshold values in Eq. (10) for  $n = 1$ ,  $n = 2$  and  $n = 3$ . The index  $l$  runs from 0 to  $n^2$  and  $K_m = k - m$  is a short hand notation.

In the case of the lowest moment  $Q_1$ , one obtains

$$Q_1(\rho) = 1 + e^{-\sigma^2} \sum_{i_1 \neq i_2} \rho_{i_1 i_2} \prod_{j=1}^d e^{in_j \mu_j}, \quad (11)$$

with  $n_j \in \{-1, 0, 1\}$ . That is, the width  $\sigma$  of the distribution enters only as multiplicative factor  $e^{-\sigma^2}$ . Since the threshold values for  $k$ -coherence and the values of  $Q_1$  for any quantum state scale with  $\sigma$  in exactly the same fashion, the question of whether  $Q_1(\rho)$  exceeds a threshold value is independent of the width of the utilized distribution  $F$ . For  $n = 1$  there is thus no advantage of generalized moments over uniform moments. As we will see in the following, however, the former are clearly advantageous for higher moments  $Q_n$  with  $n > 1$ .

### III. NUMERICAL ANALYSIS

Having established the rigorous properties of the generalized moments  $Q_n$ , it remains to identify the range of optimal values of  $\sigma$ . While narrow distributions require an accurate identification of the maximum in the interference pattern, they yield strong criteria if this maximum is found. On the other hand wider distributions give robust but potentially weaker criteria.

In the numerical studies, we thus address the following three questions:

- how many states are detected as  $k$ -coherent by the  $Q_n$ ?
- how sensitive are the  $Q_n$  with respect to mis-estimating the maxima of the interference pattern?
- how does the sensitivity of the  $Q_n$  depend on the width  $\sigma$  of the distribution  $F$ ?

To address these questions, we will consider ensembles of density matrices and determine how many states are detected to be  $k$ -coherent. We will characterize the performance of the  $Q_n$  in terms of the *detection ratio*  $R$  defined as the ratio number of states detected as  $k$ -coherent

and the ensemble size. Since there is no reliable construction of  $k$ -coherent states any ensemble will always contain states that are *not*  $k$ -coherent; the maximally achievable detection ration can therefore be substantially smaller than 1, so that its absolute value is *not* a good indicator of the strength of  $Q_n$ . To obtain an estimate of the overall strength, we will thus first compare the  $Q_n$  with previously known tools for the characterization of  $k$ -coherence and subsequently investigate the dependence of the detection ratio on errors in estimating the maximum of the interference pattern and the width  $\sigma$ .

#### A. Comparison with other criteria

In order to estimate the strength of the generalized moments  $Q_n$  we employ a comparison with a hierarchy of separability criteria that detect  $k$ -coherence [17, 24]. We shall perform this comparison based on an ensemble of 1000 states that are stationary solutions of a driven, dissipative disordered network [24].

	$k = 2$	$k = 3$	$k = 4$	$k = 5$
hierarchy	1000	989	562	12
moments	1000	969	347	0
number of states for which:				
	$k_m > k_h$	$k_m < k_h$	$k_m = k_h$	
	83	318	599	

Table II. Comparison between the hierarchy of separability criteria and the generalized moments with an asymptotically peaked distribution (*i.e.*  $\sigma \rightarrow 0$ ). Upper part: Number of states detected as  $k$ -coherent for  $k = 2, 3, 4, 5$ . Lower part: Number of states for which the generalized moments detected strictly larger, smaller and the same  $k$  as the hierarchy.

In Table II (upper part) we list the number of states that have been confirmed to be  $k$ -coherent for  $k = 2, 3, 4, 5$ , with the help of the hierarchy and the generalized moments  $Q_n$  with  $\sigma \rightarrow 0$  respectively. With either method all states are identified as 2-coherent; since only diagonal states are incoherent, the identification of 2-coherence is not challenging and it comes at no surprise that both methods perform well. In general, the moments perform particularly well in detecting low  $k$ -coherence, while for  $k = 4, 5$  the superiority of the hierarchy becomes more evident. A state-wise comparison (lower part of Table II), however, reveals that there is also a significant number of states (83 out of 1000) for which the moments detected larger  $k$  than the hierarchy. Most of these states have been identified as 4-coherent by the moments and 3-coherent by the hierarchy. Thus, even though the hierarchy is overall slightly stronger than the generalized moments, the results suggest that the performance of the moments is at least comparable to those of other known tools. This is striking since prior tools are designed to employ the information content of the full density matrix whereas the generalized moments are based on limited, easily accessible information.

## B. Dependence on distribution and errors

Having verified the overall strength of the  $Q_n$ , we are in the position to investigate how this strength depends on the width  $\sigma$  of the distribution  $f$  and errors in placing the centers of the distributions. To quantify the error in determination of  $\phi^{\max}$  providing the maximum of the interference pattern, we assume that the centers  $\mu$  of the distributions are shifted away from the maximum by some vector  $\delta$ . In our simulations we can numerically find  $\phi^{\max}$  (in the first considered ensemble of states we have by construction  $\phi^{\max} = 0$ ), randomly draw  $\delta$  and in the last step set  $\mu = \phi^{\max} + \delta$ .

The vector  $\delta$  emulates the general effect of inaccurate determination of the maximum. The latter, however, does not necessarily need to be caused by optimization issues; since the maximum of a measured interference pattern can never be determined with a precision better than that allowed by the coarse graining (binning) of the measurement, a finite  $\delta$  can also arise from experimental limitations.

If all the  $\delta_i$  vanish, the distribution is centered around the maximum of the interference pattern, and since the entire interference pattern is invariant under a global phase shift, this holds also if all the  $\delta_i$  coincide. We therefore define the invariant deviation vector  $\tilde{\delta}_i = \delta_i - \bar{\delta}$ , where the term  $\bar{\delta} = \sum_{i=1}^d \delta_i / d$  removes the aforementioned ambiguity. All the  $\tilde{\delta}_i$  are treated as normally-distributed independent random variables

$$\tilde{\delta}_i \sim \mathcal{N}(0, \sigma_G^2), \quad (12)$$

characterized by a single width  $\sigma_G$ . The particular choice  $\sigma_G = 0$  refers to the optimal case when the maximum of the pattern is identified perfectly, while positive widths cause random, but statistically controlled shifts. In order to obtain the desired width in Eq.(12), the primary, non-invariant parameters  $\delta_i$  need to be distributed according to  $\mathcal{N}(0, \frac{d}{d-1} \sigma_G^2)$ . In other words, the global phase invariance provides the effective decrease of the width by the factor  $\sqrt{(d-1)/d}$ , so that in the limit of large  $d$  the invariance in question does not result in observable effects.

Since for a fixed choice of  $\sigma_G$ , the ability to identify coherence is expected to depend on the specific choice of the  $\delta_i$ , we average the detection ratio over a large number, say  $N = 1000$ , of deviation vectors  $\delta$ , and the averaged detection ratio is denoted further by  $\langle R \rangle$ .

### 1. A single-parameter family of states

Let us start with a simple ensemble

$$\rho_a = a |\Psi_W\rangle \langle \Psi_W| + \frac{1-a}{d} \mathbf{1}, \quad (13)$$

with  $|\Psi_W\rangle = \sum_{i=1}^d |i\rangle / \sqrt{d}$  and the identity matrix  $\mathbf{1}$ , parametrized by the real parameter  $0 \leq a \leq 1$ . All  $d$  ba-

sis states are populated with equal weight independently of  $a$ ; for  $a = 1$ ,  $\rho_a$  describes a perfectly coherent superposition, and  $a = 0$  corresponds to the situation with no phase coherence. Here, we consider  $d = 7$  and try to detect  $k$ -coherence for  $k = 7$ . For both parameters  $\sigma$  and  $\sigma_G$  being fixed we numerically integrate over the range of the parameter  $a$  in order to obtain the detection ratio  $R$ . For any combination  $(\sigma, \sigma_G)$  we further average  $R$  over 1000 deviation vectors  $\delta$ . The averaged detection ratio  $\langle R \rangle$  for the third moment  $Q_3$  as a function of  $\sigma$  and  $\sigma_G$  is shown in Fig. 3.

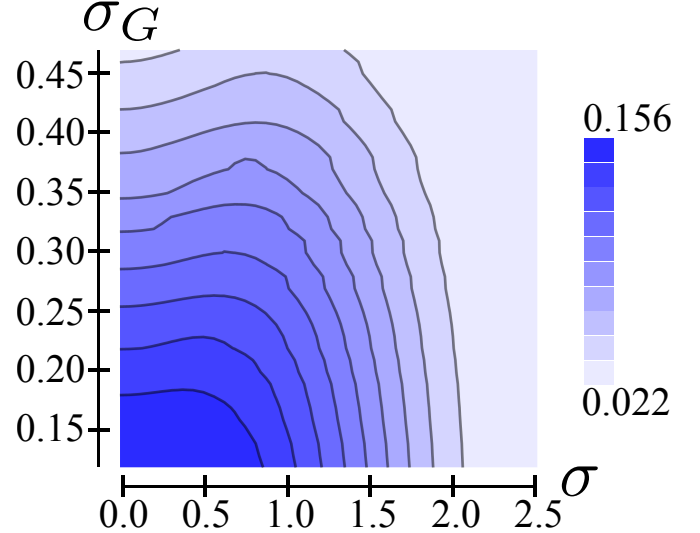


Figure 3. Dependence of the average detection ratio  $\langle R \rangle$  on width  $\sigma$  and error  $\sigma_G$  (indicated by the colorscale). The depicted data refer to  $Q_3$ ,  $d = k = 7$  and the ensemble given by Eq. (13). There is a broad region ( $\sigma_G \lesssim 0.15$  and  $\sigma \lesssim 0.7$ ) in which  $Q_3$  performs nearly as well as in the ideal case  $\sigma_G = 0$  and  $\sigma \rightarrow 0$ .

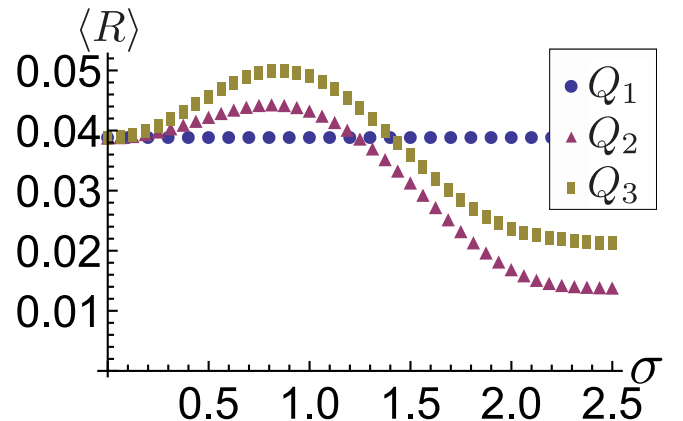


Figure 4. Dependence of the average detection ratio  $\langle R \rangle$  on width  $\sigma$  for  $\sigma_G = 0.4$ . As discussed above, the performance of  $Q_1$  is independent of  $\sigma$ , but there is an optimal, finite  $\sigma$  for  $Q_2$  and  $Q_3$ .

As expected, when  $\sigma$  is fixed,  $\langle R \rangle$  decays monotonically

cally with  $\sigma_G$ . An interesting feature can be captured while looking at the perpendicular direction, *i.e.* on the average detection ratio as a function of  $\sigma$  with fixed  $\sigma_G$ . There is a broad range of  $\sigma_G$ , for which  $\langle R \rangle$  is an increasing function of  $\sigma$ . In other words, broader distributions used during the construction of the generalized moments can compensate the inaccuracy in determination of the maximum. Fig. 4 explicitly shows the above effect for an exemplary intersection taken at  $\sigma_G = 0.4$ . The largest value of  $\langle R \rangle$  is in this case not achieved for  $\sigma = 0$ , but for a finite width  $\sigma \approx 0.9$ . In more detail, for  $\sigma = 0$  (exact evaluation of the supposed maximum of the interference pattern) the average detection ratio for  $\sigma_G = 0.4$  is 0.039 while it equals 0.166 when  $\sigma_G = 0$  (supposed maximum is the true maximum). This observation indicates that the moments with narrower distributions are rather sensitive to the error in determination of the maximum. In the suboptimal case the average detection ratio drops to  $\approx 23\%$  of the optimal value achieved for  $\sigma_G = 0$ . The larger width  $\sigma = 0.9$  (still for  $\sigma_G = 0.4$ ) provides an increase of  $\langle R \rangle$  by a factor of 1.28, to the value of 0.050.

## 2. Random states

In order to test the universal validity of the above observations we performed a similar computation with arbitrary random states  $\rho = U\Lambda U^\dagger$ , where  $\Lambda$  is a diagonal matrix describing the spectrum of  $\rho$  while  $U$  is a unitary transformation. The matrix  $U$  is drawn from the Circular Unitary Ensemble (CUE [22, 23]), while  $\Lambda$  contains squared absolute values of components from a single column of a unitary matrix, which was also generated with the help of CUE. The ensemble used to obtain the detection ratio contained 500 random states. Due to substantial computational effort we averaged the detection ratio (with fixed  $\sigma$  and  $\sigma_G$ ) over 25 deviation vectors  $\delta$ .

Let us denote by  $\langle R \rangle_{\text{ref}}$  the average detection ratio corresponding to the first moment. By construction, the same value of  $\langle R \rangle$  is obtained while using any other moment with  $\sigma = 0$ . As discussed in the context of Eq. (11), the performance of the first generalized moment is independent of the parameter  $\sigma$ . We shall thus utilize  $\langle R \rangle_{\text{ref}}$  as the reference quantity, and calculate

$$r_n = \langle R \rangle_n^{\text{max}} / \langle R \rangle_{\text{ref}}, \quad (14)$$

where by  $\langle R \rangle_n^{\text{max}}$  we denote the average detection ratio for the  $n$ th moment calculated with  $\sigma = \sigma^{\text{max}}$  being the value of  $\sigma$ , for which  $\langle R \rangle$  based on the  $n$ th moment is maximal. Finally, by  $\langle R \rangle_{\text{opt}}$  we denote the average detection ratio (independent of  $n$ ) obtained in the optimal setting  $\sigma_G = 0 = \sigma$ .

The results of our numerical analysis are presented in Table III. Comparing  $\langle R \rangle_{\text{opt}}$  with the ratios  $\langle R \rangle_{\text{ref}}$  in the suboptimal setting (positive  $\sigma_G$ ), we confirm the previous observation that the performance of the moments (those with narrow distributions) can be highly sensitive

$k$	$\langle R \rangle_{\text{opt}}$	$\sigma_G$	$\langle R \rangle_{\text{ref}}$	$r_2$	$r_3$	$\sigma^{\text{max}}$
$\geq 4$	0.55	0.3	0.45	1.00	1.01	0.3
		0.8	0.11	1.10	1.21	0.7
$\geq 5$	0.19	0.3	0.10	1.03	1.05	0.5
		0.8	0.007	1.21	1.48	0.8
$\geq 6$	0.01	0.3	0.002	1.13	1.21	0.6

Table III. Optimal (opt), reference (ref) and maximal ( $r_n$ , with respect to the reference value) average detection ratios for the first three generalized moments.  $\sigma^{\text{max}}$  is the value of  $\sigma$  for which  $\langle R \rangle$  for both  $n = 2, 3$  becomes maximal. Several values of  $k$  and  $\sigma_G$  have been tested.

to uncertainty of  $\phi^{\text{max}}$ . Similarly to the case of the first ensemble investigated, the second and third moment with positive width  $\sigma$  can significantly increase the detection ratio in this suboptimal setting.

## IV. ANALYSIS FOR MIXED STATES

In Section II we showed that the generalized moments  $Q_n$  are convex with respect to the density matrix, what further implies that their maxima are provided by pure states (see Eq. (7)). The identification of the maximum value of  $Q_n$  that can be taken for mixed states with given purity, thus allows to strengthen the detection of coherence in mixed states. As we explicitly demonstrate here for the case  $n = 2$ , there are however strongly mixed states that yield values of  $Q_2$  close to the achievable maximum. That is, the present approach is by no means limited to pure or weakly mixed states, but, it can indeed identify  $k$ -coherence also for substantially mixed states.

In Appendix B we show that the global maximum

$$\max_{\rho} (Q_2(\rho) \mid \text{Tr} \rho^2 = \mathcal{P}) \quad (15)$$

taken over all states with given purity  $\mathcal{P}$  is obtained for the state

$$\varrho_{\text{max}}(\mathcal{P}) = \left(1 - \sqrt{\frac{\mathcal{P}d - 1}{d - 1}}\right) \frac{\mathbb{1}_d}{d} + \sqrt{\frac{\mathcal{P}d - 1}{d - 1}} |W_d\rangle \langle W_d|, \quad (16)$$

with  $|W_d\rangle$  defined in Eq. (7). Moreover, any state of the form Eq. (16) with purity

$$\text{Tr} \rho^2 \leq \mathcal{P}_k = \frac{k^2 - 2k + d}{d(d - 1)}, \quad (17)$$

is at most  $k$ -coherent.  $\mathcal{P}_k$  is thus the smallest purity that permits to identify  $k + 1$ -coherence with the present moments.

Fig. 5 depicts  $Q_2(\varrho_{\text{max}}(\mathcal{P}))$  as function of  $\mathcal{P}$  (solid curve), and  $\mathcal{P}_k$  for a 5-dimensional system with  $\sigma = 1$ . The symbols (triangles, rectangles, circles) depict the numerically obtained maximum  $Q_2^{(k)}$  of  $Q_2$  with the maximization performed over all at most  $k$ -coherent states with given purity (see Eq. (B9) in the Appendix); the

horizontal lines depict the corresponding values for  $\mathcal{P} = 1$  to guide the eye.

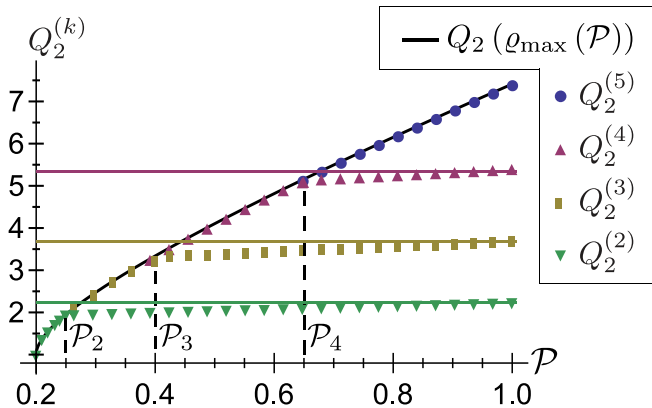


Figure 5. The values  $Q_2^{(k)}$  depict the numerically obtained, purity-dependent threshold values of  $k$ -coherence for  $k = 2, 3, 4, 5$  in a  $d = 5$ -dimensional system and  $\sigma = 1$ . The solid curve denotes the analytical result obtained for the state (16). The horizontal lines denote the original threshold values obtained by pure states.

As one can see, the values  $Q_2^{(k)}$  coincide with  $Q_2(\varrho_{\max}(\mathcal{P}))$  for  $\mathcal{P} \leq \mathcal{P}_k$ ; for  $\mathcal{P} > \mathcal{P}_k$ ,  $Q_2^{(k)}$  is nearly constant, *i.e.* there is a very small increase with  $\mathcal{P}$ . This means, that there are rather highly mixed  $k + 1$ -coherent states that yield values above the threshold values of  $Q_2^{(k)}$  for  $k$ -coherence. The approach developed hitherto is thus able to identify coherence reliably even for rather strongly mixed states, even if no information on purity is available. Since the range over which  $Q_2^{(k)}$  is nearly constant, is the larger (*i.e.* including lower values of  $\mathcal{P}$ ), the smaller  $k$  is, this holds in particular for the identification of low  $k$  coherence. That is, in particular for  $k \ll d$ , (as it typically is the case in excitation transport [27]), the present framework can detect coherence very well even for quantum states with substantial degree of mixing; but if necessary, one may always resort to the purity-dependent threshold values in order to improve the detection.

## V. CONCLUSION

As we have seen, an interference pattern permits to draw rigorous conclusions on coherence beyond the intuitive, qualitative expectation that an interference contrast grows with increasing coherence properties. From a practical point of view, the freedom in choice of sampling as well as the possibility to include additional information (like purity) makes this approach flexible, so that it can be tailored for the specific properties of a system under investigation. That is, limitations on experimentally variable quantities may be compensated through suitably chosen distributions with variable widths that reflect the realistically achievable sampling. In particular, with

noisy data that does not permit to reconstruct an entire interference pattern reliably, the generalized moments of low order still can characterize coherence properties reliably.

Here, we have been considering the case of  $d$  independently adjustable phases, but the underlying framework can be generalized also for the variation of fewer phases. Moreover, we developed the general framework which can utilize a wrapped version of an arbitrary probability distribution. While the normal distribution seems to be the most natural first choice, more sophisticated distributions can even better support a particular experimental realization. Beyond the conceptual connection between a directly observable interference pattern and the underlying, abstract coherence properties, the present approach thus provides a versatile method to characterize coherence properties in a wide range of systems.

At the end, let us establish a general link between the generalized moments discussed in this paper and the commonly employed characterization of quantum coherence in terms of the  $l_1$ -norm of coherence  $C_{l_1}(\rho) = \sum_{i \neq j} |\rho_{ij}|$  [13]. Since the interference pattern (Eq. (1)) satisfies  $P(\phi) \leq 1 + C_{l_1}$  for any value of  $\phi$ ,  $C_{l_1}$  is bounded,  $C_{l_1}(\rho) \geq |Q_n(\rho)|^{1/n} - 1$ , in terms of the moments  $Q_n(\rho)$ . It is thus not surprising, that  $Q_n$  contains enough information to provide a valuable description of quantum coherence as well as practical criteria for identification of  $k$ -coherence.

## ACKNOWLEDGMENTS

We thank Alexander Stibor and Björn Witt for interesting and helpful conversations. Financial support by the European Research Council under the project Odyquant is gratefully acknowledged.

## Appendix A: Derivation of threshold values

We start the derivation by inserting  $\rho^{(k)} = |\Psi^{(k)}\rangle\langle\Psi^{(k)}|$  with  $|\Psi^{(k)}\rangle$  given by (9) into the expression (5):

$$\begin{aligned} Q_n &= \sum_{i_1, i_2, \dots, i_{2n}=1}^d \sqrt{\lambda_{i_1} \lambda_{i_2} \dots \lambda_{i_{2n}}} \prod_{i=1}^d R_{n_i} \cos X(\boldsymbol{\mu}) \\ &\leq \sum_{i_1, i_2, \dots, i_{2n}=1}^d \sqrt{\lambda_{i_1} \lambda_{i_2} \dots \lambda_{i_{2n}}} \prod_{j=1}^d R_{n_j} \\ &\equiv g_n(\boldsymbol{\lambda}), \end{aligned} \quad (\text{A1})$$

with

$$X(\boldsymbol{\mu}) = \sum_{j=1}^d n_j (\mu_j - \varphi_j). \quad (\text{A2})$$

An estimate  $\cos(\cdot) \leq 1$  applied in the second line implies that the maximum of  $Q_n$  with respect to  $\boldsymbol{\mu}$  is achieved



when the peak position of the probability distribution coincides with the maximum of the interference pattern.

In the next step we employ the concept of Schur-concavity [25, 26]. For any two vectors  $\lambda$  and  $\lambda'$  such that  $\lambda'$  is majorized by  $\lambda$  (so that  $\lambda' \prec \lambda$ ) and any Schur-concave function  $g(\lambda)$  one gets  $g(\lambda') \geq g(\lambda)$ . In the case of pure,  $k$ -coherent states all vectors  $\lambda$  majorize the uniform vector

$$\lambda_k = \left( \underbrace{\frac{1}{k}, \frac{1}{k}, \dots, \frac{1}{k}}_{k\text{-times}}, \underbrace{0, \dots, 0}_{(d-k)\text{-times}} \right). \quad (\text{A3})$$

To finish the proof we thus only need to show that the function  $g_n(\lambda)$  defined in (A1) is Schur-concave for  $n = 2$  and  $n = 3$ . To this end it is sufficient to show that  $g_n(\lambda)$

satisfies the condition

$$S_{ij}(g) = (\lambda_i - \lambda_j) \left( \frac{\partial g}{\partial \lambda_i} - \frac{\partial g}{\partial \lambda_j} \right) \leq 0 \quad \forall i, j = 1, 2, \dots, d. \quad (\text{A4})$$

To proceed further we need an explicit form of both functions; that is why we define

$$W_A^B = \prod_{l=1}^A \lambda_{i_l} \prod_{n=1}^B \sqrt{\lambda_{j_n}}, \quad (\text{A5})$$

and

$$G_{AB} = \sum_{\neq} W_A^B, \quad (\text{A6})$$

where  $\sum_{\neq}$  denotes the sum over all pairwise different indices  $i_1, \dots, i_A, j_1, \dots, j_B$  running from 1 to  $d$ . With this, one obtains

$$g_2(\lambda) = 1 + (1 + R_2^2) G_{20} + 2R_1^2 G_{02} + 2R_1^2 (1 + R_2) G_{12} + R_1^4 G_{04}, \quad (\text{A7})$$

and

$$\begin{aligned} g_3(\lambda) = & 1 + 3R_1^2 G_{02} + 3(1 + R_2^2) G_{20} + 6R_1^2 (1 + R_2) G_{12} + 3R_1^4 G_{04} + 2(1 + 3R_2^2) G_{30} + 3R_1^2 (3 + 4R_2 + 3R_2^2) G_{22} \\ & + 6R_1^4 (1 + R_2) G_{14} + R_1^6 G_{06} + 6R_1 (2R_1 + R_1 R_2 + R_2 R_3) \sum_{\neq} \sqrt{\lambda_i}^3 \lambda_j \sqrt{\lambda_m} \\ & + 2R_1^3 (3R_1 + R_3) \sum_{\neq} \sqrt{\lambda_i}^3 \sqrt{\lambda_j \lambda_m \lambda_l} + (3R_1^2 + R_3^2) \sum_{\neq} \sqrt{\lambda_i \lambda_j}^3. \end{aligned} \quad (\text{A8})$$

Since all parameters  $R_n$  are non-negative we can treat each term in (A7-A8) separately and find (for  $i \neq j$ )

$$\begin{aligned} \frac{\partial G_{AB}}{\partial \lambda_i} = & A \sum_{\neq, \setminus \{i\}} W_{A-1}^B + \frac{B}{2\sqrt{\lambda_i}} \sum_{\neq, \setminus \{i\}} W_A^{B-1} \\ = & A \sum_{\neq, \setminus \{i, j\}} \left( W_{A-1}^B + (A-1)\lambda_j W_{A-2}^B + B\sqrt{\lambda_j} W_{A-1}^{B-1} \right) \\ & + \frac{B}{2\sqrt{\lambda_i \lambda_j}} \sum_{\neq, \setminus \{i, j\}} \left( \sqrt{\lambda_j} W_A^{B-1} + A\lambda_j^{3/2} W_{A-1}^{B-1} + (B-1)\lambda_j W_A^{B-2} \right), \end{aligned} \quad (\text{A9})$$

where  $\sum_{\neq, \setminus Z}$  denotes the sum  $\sum_{\neq}$  with the additional exclusion of the values  $Z$ . From Eq. (A10) one may see that  $S_{ij}(G_{AB})$  (with  $i \neq j$ ) contains only terms proportional to  $-(\lambda_i - \lambda_j)^2$ ,  $(\lambda_i - \lambda_j)(\sqrt{\lambda_j} - \sqrt{\lambda_i})$  and  $(\lambda_i - \lambda_j)(\lambda_j^{3/2} - \lambda_i^{3/2})$ . Since all these terms are separately non-positive,  $g_2(\lambda)$  is a sum of Schur-concave functions  $G_{AB}$ , and thus Schur concave.

The function  $g_3(\lambda)$  involves additional terms which are not of the form of  $G_{AB}$  and are not Schur-concave. It is,

however, possible to show that the terms

$$2R_1^3 (3R_1 + R_3) \sum_{\neq} \sqrt{\lambda_i}^3 \sqrt{\lambda_j \lambda_m \lambda_l} + 6R_1^4 (1 + R_2) G_{14}, \quad (\text{A11})$$

$$\text{and } 6R_1 (2R_1 + R_1 R_2 + R_2 R_3) \sum_{\neq} \sqrt{\lambda_i}^3 \lambda_j \sqrt{\lambda_m} +$$

$$(3R_1^2 + R_3^2) \sum_{\neq} \sqrt{\lambda_i \lambda_j}^3 + 3R_1^2 (3 + 4R_2 + 3R_2^2) G_{22},$$

are Schur-concave for  $R_1 \geq R_3$ . For the wrapped normal distribution we get  $R_1 = e^{-\sigma^2/2} \geq e^{-9\sigma^2/2} = R_3$ , so that  $g_3(\lambda)$  is Schur-concave as well.



## Appendix B: Derivation of the global maximum

Since the generalized moments (5) are real and non-negative we have the estimate

$$Q_2 = |Q_2| \leq \sum_{i_1, i_2, i_3, i_4=1}^d |\rho_{i_1 i_3}| |\rho_{i_2 i_4}| \prod_{j=1}^d R_{n_j}. \quad (\text{B1})$$

Since the right hand side does not depend on  $\mu$ , the same upper bound applies to the maximum of  $|Q_2|$  (maximized with respect to the center of the distribution  $F(\phi)$ ).

The right hand side of (B1) explicitly reads

$$1 + (1 + R_2^2) \sum_{\neq} |\rho_{ij}|^2 + 2R_1^2 \sum_{\neq} |\rho_{ij}| + R_1^4 \sum_{\neq} |\rho_{ij}| |\rho_{kl}| + 2R_1^2 (1 + R_2) \sum_{\neq} |\rho_{im}| |\rho_{ji}|, \quad (\text{B2})$$

and saturates if  $\mathbb{R} \ni \rho_{ij} \geq 0$  for all  $i \neq j$ , so that the density matrix has only real and non-negative entries.

Using the inequality of arithmetic and geometric mean we get the estimates

$$\sum_{\neq} |\rho_{im}| |\rho_{ji}| \leq (d-2) \sum_{\neq} |\rho_{ij}|^2, \quad (\text{B3})$$

$$\sum_{\neq} |\rho_{ij}| |\rho_{ml}| \leq (d-3)(d-2) \sum_{\neq} |\rho_{ij}|^2. \quad (\text{B4})$$

Since we assume that the purity  $\mathcal{P} = \text{Tr} \rho^2$  is fixed we get

$$\sum_{\neq} |\rho_{ij}|^2 = \mathcal{P} - \sum_{i=1}^d \rho_{ii}^2, \quad (\text{B5})$$

what also implies that

$$\sum_{\neq} |\rho_{ij}| \leq \sqrt{d(d-1) \left( \mathcal{P} - \sum_{i=1}^d \rho_{ii}^2 \right)}. \quad (\text{B6})$$

The maximum of (B5), as well as maxima of (B3, B4) and (B6) are provided by the uniform distribution

$$\forall_i \rho_{ii} = \frac{1}{d}. \quad (\text{B7})$$

This observation immediately leads to a state independent maximum of  $Q_2$ . The above maximum may be saturated only when the inequalities used in (B3, B4) and

(B6) saturate too, i.e. when

$$|\rho_{ij}| = \frac{1}{d} \sqrt{\frac{\mathcal{P}d-1}{(d-1)}}, \quad \forall_{i \neq j}. \quad (\text{B8})$$

The last conclusion proves Eq. (16), showing that the maximum is always global.

The next step is to determine when the state (16) does not happen to be  $k+1$ -coherent. For example, if  $\mathcal{P} = 1$ , the global maximum is attained by a  $d$ -coherent pure state. We start with the following observation: to realize any  $k$ -coherent state it is sufficient to consider the form:

$$\rho^{(k)} = \sum_{m=1}^{D_k} \sum_{i,j \in I_m(k)} \Xi_{ij}^{(m)} |i\rangle \langle j|, \quad D_k = \binom{d}{k}, \quad (\text{B9})$$

with  $I_m(k)$  being for each  $m = 1, \dots, D_k$  a unique set of  $k$  different indices taken from  $\{1, \dots, d\}$ . By  $\Xi_{ij}^{(m)}$  we denote arbitrary (not normalized)  $k \times k$  positive semi-definite matrices.

We shall now construct  $\rho^{(k)}$  with all diagonal elements equal to  $\kappa \equiv 1/d$  and all off-diagonal elements equal to  $0 \leq \beta \leq 1/d$ . The maximal value of  $\beta$  is provided by the case when for every  $m$ , all diagonal elements of  $\Xi_{ij}^{(m)}$  are equal to some  $h > 0$  and all corresponding off-diagonal elements are given by some  $b \leq h$ . From combinatorial considerations we find

$$\frac{1}{d} \equiv \kappa = \binom{d-1}{k-1} h, \quad \beta = \binom{d-2}{k-2} b. \quad (\text{B10})$$

We thus obtain

$$\begin{aligned} \text{Tr} \left( \rho^{(k)} \right)^2 &= \frac{1}{d} + d(d-1) \beta^2 \\ &= \frac{1}{d} + d(d-1) \left[ \binom{d-2}{k-2} \right]^2 b^2 \\ &\leq \frac{1}{d} + d(d-1) \left[ \binom{d-2}{k-2} \right]^2 h^2 \\ &= \mathcal{P}_k. \end{aligned} \quad (\text{B11})$$

In that way we have recovered the range given in Eq. (17).

## REFERENCES

- 
- [1] C. K. Hong, Z. Y. Ou, and L. Mandel, Phys. Rev. Lett. **59**, 2044 (1987).
  - [2] O. Nairz, M. Arndt, and A. Zeilinger, Am. J. Phys. **71**, 319 (2003).
  - [3] P. Facchi, J. Mod. Opt. **51**, 1049 (2004).
  - [4] S. Eibenberger, X. Cheng, J. P. Cotter, M. and Arndt, Phys. Rev. Lett. **112**, 250402 (2014).
  - [5] F. Fassioli, R. Dinshaw, P. C. Arpin, and G. D. Scholes,

- J. R. Soc. Interface **11**, 20130901 (2014).
- [6] T. Scholak, F. de Melo, T. Wellens, F. Mintert, and A. Buchleitner, Phys. Rev. E **83**, 021912 (2011).
  - [7] M. O. Scully, M. S. Zubairy, G. S. Agarwal and H. Walther, Science **299**, 862 (2003).
  - [8] M. Lostaglio, K. Korzekwa, D. Jennings, and T. Rudolph, Phys. Rev. X **5**, 021001 (2015).
  - [9] D. Girolami, Phys. Rev. Lett. **113**, 170401 (2014).
  - [10] F. Levi and F. Mintert, New J. Phys. **16**, 033007 (2014).
  - [11] C. Smyth and G. D. Scholes, Phys. Rev. A **90**, 032312 (2014).
  - [12] R. Horodecki et al, Rev. Mod. Phys. **81**, 865 (2009).
  - [13] M. Cramer, T. Baumgratz, and M. Plenio, Phys. Rev. Lett. **113**, 140401 (2014).
  - [14] J. Åberg, Phys. Rev. Lett. **113**, 150402 (2014).
  - [15] V. Vedral, M. Arndt, and T. Juffmann, HFSP Journal **3**, 386 (2009).
  - [16] T. R. Bromley, M. Cianciaruso, and G. Adesso, Phys. Rev. Lett. **114**, 210401 (2015).
  - [17] F. Levi and F. Mintert, Phys. Rev. Lett. **110**, 150402 (2013).
  - [18] A. Streltsov, U. Singh, H. Shekhar Dhar, M. Nath Bera and G. Adesso, Phys. Rev. Lett. **115**, 020403 (2015).
  - [19] D. P. Pires, L. C. Céleri, and D. O. Soares-Pinto, Phys. Rev. A **91**, 042330 (2015).
  - [20] J. Flusser, B. Zitova, T. Suk, *Moments and Moment Invariants in Pattern Recognition*, Wiley & Sons Ltd., 2009, sec. 6.5, p. 204
  - [21] S. R. Jammalamadaka and A. SenGupta, *Topics in circular statistics* (World Scientific), 2001
  - [22] K. Życzkowski and M. Kuś, J. Phys. A: Math. Gen. **27**, 4235 (1994).
  - [23] K. Życzkowski and H-J. Sommers, J. Phys. A: Math. Gen. **34**, 7111 (2001).
  - [24] B. Witt and F. Mintert, New J. Phys. **15**, 093020 (2013).
  - [25] S. Wang, T. Zhang, and B. Xi, *Schur convexity for a class of symmetric functions. In Information Computing and Applications* (Springer Berlin Heidelberg), 2011
  - [26] A. W. Marshall and I. Olkin, *Inequalities: Theory of Majorization and Its Applications* (Academic Press), 1979
  - [27] F. Levi, S. Mostarda, F. Rao and F. Mintert, Rep. Prog. Phys. **78**, 082001 (2015).